REMARKS

The following remarks are fully and completely responsive to the Office Action dated June 20, 2002. Claims 1-18 are pending in this application. In the outstanding Office Action, the drawings were objected to; the title of the invention was objected to; the disclosure was objected to; claims 2, 4, 5 and 6 were objected to; and claims 1-18 were rejected under 35 U.S.C. § 103(a) (four different rejections). No new matter has been entered. Claims 1-18 are presented for consideration.

Objection to the Drawings

The drawings were objected to because there was no table 6g in Figure 8. Applicant has amended the specification to remove the reference numeral "6g". Accordingly, Applicant respectfully requests reconsideration and withdrawal of this objection to the drawings.

The drawings were also objected to as failing to comply with 37 C.F.R. § 1.84(p)(5) because the reference sign 210 of Figure 10 is not mentioned in the description. Applicant has amended the specification at page 30 to include the reference numeral 210. Accordingly, Applicant respectfully asserts that the drawings comply with 37 C.F.R. § 1.84(p)(5) and requests reconsideration and withdrawal of this objection to the drawings.

Objection to the Specification

The disclosure was objected to because there is no table 6g in Figure 8. The disclosure was also objected to because the word "large" should be "loud". Applicant's

amendments to the specification have removed the reference numeral "6g" from the specification and Applicant's amendment to the specification has replaced the word "large" with the word "loud" on page 2 at lines 20 and 24 as suggested by the Examiner. Accordingly, Applicant respectfully requests reconsideration and withdrawal of the objection to the specification.

Objection to the Title

Applicant has amended the title to a new title that is clearly indicative of the invention to which the claims are directed. In amending the title, Applicant has utilized the title suggested by the Examiner. Accordingly, Applicant respectfully requests reconsideration and withdrawal of the objection to the title.

Claim Objections

Claims 2, 4, 5 and 6 were objected to because of informalities recited in the Office Action dated June 20, 2002. The amendments to claims 2, 4, 5 and 6 correct these informalities. Accordingly, Applicant respectfully requests reconsideration and withdrawal of the objection to claims 2, 4, 5 and 6.

35 U.S.C. § 103(a)

Claims 1, 7 and 13 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Everhart (U.S. Patent No. 6,230,138). In making this rejection, the Office Action asserts that this reference teaches and/or suggests each and every

(09/651,058)

element of the claimed invention. Applicant respectfully requests reconsideration of this rejection.

Claim 1 recites a speech recognition system. The system includes a plurality of voice pickup means for picking up uttered voices. Determination means determines a speech signal suitable for speech recognition from speech signals output from the plurality of voice pickup means. Speech recognition means performs speech recognition based on the speech signal determined by the determination means.

Claim 7 recites a speech recognition system that includes a plurality of voice pickup sections for picking up uttered voices. A determination section determines a speech signal suitable for speech recognition from speech signals output from the plurality of voice pickup sections. A speech recognizer performs speech recognition based on the speech signal determined by the determination section.

Claim 13 recites a speech recognition method for a speech recognition system having a plurality of voice pickup means for picking up voices. This method includes a voice pickup step of picking up uttered voices using the plurality of voice pickup means. A determination step determines a speech signal suitable for speech recognition from speech signals output from the plurality of voice pickup means. A speech recognition step performs speech recognition based on the speech signal determined by the determination step.

Everhart discloses a method and apparatus for controlling multiple speech engines in an in-vehicle speech recognition system. This system includes a microphone 42 and pushed-to-talk (PTT) control 44 mounted near each speaking location. Once a driver/passenger depresses a PTT control 44, a selector 76 receives a

(09/651,058)

TECH/136295.1

PTT control signal. In some embodiments, each PTT control 44 produces a unique location signal so that the selector 76 can distinguish the control signals into an address. The main processor 48 of the selector 76 then processes the location signal using a selection application 56 to determine which speech engine recognition algorithm should be used to process the audio signal from the microphone 42.

In Everhart, the selector 76 then relays a PTT selection signal to the selected speech engine 74. Thereafter, a listening mode within the selected speech processor 70 is initiated and the speech processor 70 receives the audio signals from the selected microphone 42. The speech processor then processes the digitized signals using the recognition algorithms of the selected speech engine to identify a matching voice command from the selected active grammar.

Accordingly, it appears that Everhart teaches a plurality of voice pickup means for picking up voices and speech recognition means for performing speech recognition.

The Office Action asserts that determining the speech signal suitable for speech recognition from speech signals output from the plurality of voice pickup means is a well-known part of a speech recognition engine. Therefore, the Office Action asserts that it would be obvious for a person with ordinary skill in the art of speech signal processing to include a determining means for determining speech signal suitable for speech recognition from speech signals output from said plurality of voice pickup means to avoid errors produced by miss-recognizing noise as voice.

The Office Action has not provided any reference illustrating that this concept is a well-known part of any speech recognition engine. Accordingly, Applicant respectfully requests the Examiner provide a reference that teaches this concept. Furthermore, it

(09/651,058)

TECH/136295.1

appears that Everhart teaches away from using the recited determining means since Everhart utilizes a push-to-talk control 44 to identify the location of the microphone which is receiving the voice command. Accordingly, it does not appear obvious that one of ordinary skill in the art would add the recited determining means to Everhart since Everhart uses a push-to-talk control 44 to identify which microphone should be active and monitored for a voice command.

Everhart fails to teach and/or suggest the claimed invention. Specifically, Everhart fails to teach and/or suggest a determining means for determining a speech signal suitable for speech recognition from speech signals output from the plurality of voice pickup means. Consequently, Applicant requests reconsideration and withdrawal of this rejection of claims 1, 7 and 13 under 35 U.S.C. § 103(a).

Claims 2, 8 and 14 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Everhart (discussed above) in view of Fedele (U.S. Patent No. 4,627,091). In making this rejection, the Office Action asserts that the combination of these two references teaches and/or suggests each and every element of the claimed invention. The Office Action also asserts that it would be obvious to one of ordinary skill in the art to combine these two references. Applicant respectfully requests reconsideration of this rejection.

The Office Action admits that Everhart fails to teach and/or disclose any specific component or process that selects the speech signals output from the plurality of voice pickup means whose level is equal to or higher than a predetermined speech level and continues over a predetermined time as a speech signal suitable for speech recognition. The Office Action cites Fedele as correcting this deficiency in Everhart.

(09/651,058)

Fedele teaches in the Background section that it is well known in speech recognition systems to select a speech level equal to or higher than a predetermined level as a speech signal suitable for processing. Fedele also teaches a speech detecting apparatus, where not only the voice portion but also the unvoiced portion of the speech can be detected and stored. Thus, Fedele teaches holding data corresponding to a fixed time period prior to the voice speech in which the amplitude of data is larger than a predetermined value. Fedele does not disclose, however, that the speech signals output from the plurality of voice pickup means whose speech level is equal to or higher than a predetermined speech level and continues over a predetermined period of time is selected as the speech signal suitable for speech recognition.

The combination of Fedele and Everhart fails to teach and/or suggest each and every element of the claimed invention. The combination of these two references fails to teach and/or suggest the recited determination means for determining a speech signal suitable for speech recognition from speech signals output from the plurality of voice pickup means. The combination of these two references also fails to teach and/or suggest that the speech signals output from the plurality of voice pickup means whose speech level is equal to or higher than a predetermined speech level and continues over a predetermined time period is selected as the speech signal for speech recognition. Accordingly, Applicant requests reconsideration and withdrawal of the rejection of claims 2, 8 and 14 under 35 U.S.C. § 103(a).

Claims 3, 9 and 15 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Everhart (discussed above) in view of Bowen (U.S. Patent

(09/651,058)

No. 5,561,737). In making this rejection, the Office Action asserts that the combination of these two references teaches each and every element of the claimed invention. The Office Action also asserts that one of ordinary skill in the art would combine these two references.

In making this rejection, the Office Action admits that Everhart fails to teach and/or suggest that the determination means requires an average S/N value and average voice power of each of the speech signals output from the plurality of voice pickup means and determines the speech signals whose average S/N value and average voice power are greater than respective predetermined threshold values as the speech signal suitable for speech recognition. This Office Action cites Bowen as correcting this deficiency in Everhart.

Bowen discloses a system for selecting microphones for a telephone conference. In this system, the input energy of each microphone and the difference between the maximum value and the minimum value of the input energy is used to determine that the voice exists when the difference is greater than a predetermined value. The microphone having the maximum value is determined as the voice source. Thus, the microphone is selected as the voice source only by the amplitude of the voice energy.

Bowen also teaches calculating short- and long-term energy for each microphone input as a method for selecting the microphone input to make active. As illustrated in column 9, beginning at line 5, these averages are calculated using the average of the peak-absolute-value selected for each microphone input. Bowen, however, fails to teach and/or suggest using an average signal-to-noise S/N value in this calculation.

The device taught by Bowen provides a signal for selecting each one of the five input circuits for respectively providing its microphone signal to the digital signal processor 110 via five serial-to-parallel converters. Consequently, the microphone input signals are weighted and summed together by the digital signal processor 110 to form the desired unitary microphone output signal. It appears, however, that Bowen fails to teach and/or suggest selecting a single speech signal for processing. Bowen also fails to overcome the deficiencies noted above in Everhart.

The combination of Bowen and Everhart fails to teach and/or suggest each and every element of the claimed invention. The combination of these two references fails to teach and/or suggest the determination means for determining a speech signal suitable for speech recognition from speech signals output from the plurality of voice pickup means. The combination of these references also fails to teach a determination means that requires an average S/N value and average voice power of each of the speech signals output from the plurality of voice pickup means and determines the speech signal, whose average S/N value and average voice power are greater than respective predetermined threshold values, as the speech signal suitable for speech recognition. Therefore, Applicant respectfully request reconsideration and withdrawal of the rejections of claims 3, 9 and 15 under 35 U.S.C. § 103(a).

Claims 4-6, 10-12, and 16-18 were rejected under 35 U.S.C. § 103(a) as being unpatentable over the combination of Everhart, Fedele and Bowen (all discussed above). In making this rejection, the Office Action asserts that the combination of these three references teaches and/or suggests each and every element of the claimed invention. The Office Action also asserts that it would be obvious to one of ordinary skill

(09/651,058)

in the art to combine these three references. Applicant respectfully requests reconsideration and withdrawal of this rejection under 35 U.S.C. § 103(a).

In this rejection, the Office Action asserts that Bowen teaches and/or suggests processing the speech signals by order of dominance so as to determine an order of processing by determining a candidate order of those speech signals whose average S/N values and average voice powers are greater than the respective predetermined threshold values. As discussed in detail above, Bowen fails to disclose and/or suggest using the average S/N values. Bowen also fails to disclose determining an order of the candidates suitable for speech recognition. Specifically, Bowen teaches combining any speech signal whose average voice power is above a predetermined value. Accordingly, Bowen fails to teach and/or suggest selecting a single candidate for speech recognition. Furthermore, Bowen fails to teach determining an order of the candidates for speech recognition.

As also discussed above, the combination of these three references fails to teach and/or suggest each and every limitation of claims 1, 7 and 13 from which these claims depend either directly or indirectly. Accordingly, for these reasons and those discussed above, Applicant respectfully requests reconsideration and withdrawal of the rejection of claims 4-6, 10-12 and 16-18 under 35 U.S.C. § 103(a).

Conclusion

Applicant's amendments and remarks have overcome the objections and rejections set forth in the Office Action dated June 20, 2002. Specifically, Applicant's amendments to the Title have overcome the objection to the Title. Applicant's

- 13 -   (09/651,058)

amendments to the Specification overcome the objections to the Drawings. Applicant's amendments to the Specification overcome the objections to the Specification. Applicant's remarks have distinguished claims 1-18 from the cited prior art and thus overcome the four rejections of these claims under 35 U.S.C. § 103(a). Accordingly, claims 1-18 are in condition for allowance. Therefore, Applicant respectfully requests consideration and allowance of claims 1-18.

Applicant submits that the application is now in condition for allowance. If the Examiner believes the application is not in condition for allowance, Applicant respectfully requests that the Examiner contact the undersigned attorney by telephone if it is believed that such contact will expedite the prosecution of the application.

The Commissioner is authorized to charge payment for any additional fees which may be required with respect to this paper to Deposit Account No. 01-2300, making reference to attorney docket number 107156-00019.

Respectfully submitted,

Rustan J. Hill
Registration No. 37,351

Customer No. 004372
ARENT FOX KINTNER PLOTKIN & KAHN, PLLC
1050 Connecticut Avenue, N.W., Suite 400
Washington, D.C. 20036-5339
Tel: (202) 857-6000
Fax: (202) 638-4810
GEO/RJH/mzk

Enclosures: Marked-Up Copy of Amended Paragraphs in Specification
Marked-Up Copy of Amended Claims

- 14 -   (09/651,058)

MARKED-UP COPY OF AMENDED PARAGRAPHS IN THE SPECIFICATION
AS REQUIRED UNDER 37 C.F.R. § 1.121

Please replace the paragraph that begins on page 2, line 19, with the following paragraph:

Other passengers who are seated far from the microphone should therefore utter [large] loud voices toward the microphone to secure a sufficient input voice level. To improve the speech recognition precision of such a speech recognition system, other passengers than the driver should also utter [large] loud voices toward the microphone to input uttered speeches into the microphone without being affected by noise in a vehicle.

Please replace the paragraph that begins on page 3 at line 1 with the following paragraph:

Accordingly, it is an object of the present invention to provide a speech recognition system which has an improved operability and can allow more than one person to secure a sufficient input voice level without uttering [large] loud voices or without being affected by ambient noise.

Please replace the paragraph that begins on page 23 at line 14 with the following paragraph:

The individual cases 1, 2, 3 and so forth in the noise selection table [6g] shown in FIG. 8 are preset based on the results of experiments on the voice characteristics obtained when passengers actually uttered voices at various positions in a vehicle in which all the microphones M_1-M_N were actually installed.

- 15 -

(09/651,058)

Please replace the paragraph that begins on page 30 at line 12 with the following paragraph:

In the next step 208, the speech recognizer 7 read the speech frame data and noise frame data most suitable for speech recognition from the storage section 4, performs speech recognition on the read speech frame data and noise frame data, and terminates a sequence of speech recognition processes when an adequate speech recognition result is acquired as determined by step 210.

Please replace the paragraph that begins on page 30 at line 19 with the following paragraph:

When no adequate speech recognition result is acquired in step 210, on the other hand, the speech recognizer 7 checks in step 212 if there are next candidates of speech frame data and noise frame data, reads the next candidates of speech frame data and noise frame data, if present, from the storage section 4 and repeats the sequence of processes starting at step 208. When no adequate speech recognition result is obtained even after re-execution of the speech recognition, the speech recognizer 7 likewise reads next candidates of speech frame data and noise frame data from the storage section 4 and repeats the sequence of processes in steps 208 to 212 until the adequate speech recognition result is obtained.

(09/651,058)

<u>MARKED-UP COPY OF AMENDED CLAIMS</u>
<u>AS REQUIRED UNDER 37 C.F.R. § 1.121</u>

Please amend claims 2, 4, 5 and 6 as follows:

2.   (Amended) The speech recognition system according to claim 1, wherein that of said speech signals output from said plurality of voice pickup means whose speech level is equal to or higher than a predetermined speech level and continues over a predetermined period of time is [determined] <u>selected</u> as said speech signal suitable for speech recognition.

4.   (Amended) The speech recognition system according to claim 3, wherein<u>:</u>

said determination means determines a candidate order of those speech signals whose average S/N values and average voice powers are greater than said respective predetermined threshold values and which are candidates for said speech signal suitable for speech recognition, in accordance with said average S/N values and average voice powers; and

said speech recognition means sequentially executes speech recognition on said candidates in accordance with said candidate order from a highest candidate to a lower one.

5.   (Amended) The speech recognition system according to any one of claims 1 to 4, wherein said determination means [treats] <u>comprises:</u>

<u>treating</u> those of said speech signals which are other than said speech signal suitable for speech recognition as noise signals.

(09/651,058)

6. (Amended) The speech recognition system according to any one of claims 1 to [5] 4 , wherein the meaning of other speech signals than said speech signal suitable for speech recognition, is determined to be that those speech [signal] signals whose average S/N value and average voice power become minimum [is] and that such signals are treated as [a] noise [signal] by said determination means.

TECH/136295.1